20

CLAIMS:

1.   A method of searching in a collection of documents, the documents having a
tree-like structure and each document in the collection of documents complying with at least
one document structure definition in a collection of document structure definitions,
comprising the steps of:

5   -   receiving (1) a certain branch;

       -          determining (3) a subset of the collection of document structure definitions,
each document structure definition in the subset allowing the certain branch to exist in a
document complying to the document structure definition;

       -          determining (5) a subset of the collection of documents, the subset of

10   documents comprising all documents of the collection of documents complying to any one of
the document structure definitions in the subset; and

       -          searching (7) for at least part of the certain branch in each document.

2.   A method as claimed in claim 1, wherein a further step comprises attempting

15   (9) to decrypt each encrypted document in the subset of documents.

3.   A method as claimed in claim 1, wherein the step of determining a subset of
the collection of documents comprises calculating (21) a number for at least part of the
certain branch by applying a hash function to the at least part of the certain branch and

20   looking up (23) which documents are mapped to the calculated number in a mapping from
number to documents, the mapping being associated with a document structure definition of
the subset of document structure definitions and the documents in the mapping complying to
the document structure definition.

25   4.   A method as claimed in claim 3, wherein a further step comprises receiving
(27) a certain value associated with the certain branch, the mapping further comprises an
association between a document in the mapping and a value domain partition, and the step of
determining a subset of the collection of documents further comprises checking (29) whether
. a value domain partition associated to a document mapped to the calculated number matches

21

a further value domain partition, the further value domain partition comprising the received value.

5.   A method as claimed 1, wherein the step of determining a subset of the collection of documents comprises looking up (25), in a mapping from document structure definition to documents, which documents comply to any one of the subset of document structure definitions.

6.   A method as claimed in claim 1, wherein the step of determining a subset of the collection of document structure definitions comprises calculating (11) a further number for at least part of the certain branch by applying a further hash function to the at least part of the certain branch and looking up (13) which document structure definitions are mapped to the calculated number in a mapping from number to document structure definitions.

7.   A method as claimed in claim 1, wherein the step of determining a subset of the collection of document structure definitions comprises attempting (15) to decrypt each encrypted document structure definition in the collection of document structure definitions and attempting (17) to determine for each document structure definition whether the document structure definition allows the certain branch to exist in a document complying to the document structure definition

8.   A computer program product enabling a programmable device to carry out a method as claimed in claim 1.

9.   An electronic device (31) for searching in a collection of documents, comprising electronic circuitry (33), the electronic circuitry functionally comprising (33):
-       an input receiver (35) for receiving a certain branch;
-       a definition subset determiner (37) for determining a subset of a collection of document structure definitions, each document structure definition in the subset allowing the certain branch to exist in a document complying to the document structure definition;
-       a document subset determiner (39) for determining a subset of a collection of documents, the subset of documents comprising all documents of the collection of documents complying to any one of the document structure definitions in the subset; and
-       a searcher (41) for searching for at least part of the certain branch in each document.

22

10.   A method of indexing a collection of documents, the documents having a tree-like structure and each document in the collection of documents complying to at least one document structure definition in a collection of document structure definitions, comprising
5   the steps of:

-          creating (51) an empty index for each document structure definition of the collection of document structure definitions, the index mapping an integer from a range of integers to a document of the collection of documents;

-          calculating (53) a number for at least a part of a branch in a document of the
10   collection of documents by applying a hash function to the at least part of the branch, the number being limited to the range of integers and the calculation possibly producing a same number for different branches; and

-          creating (55) an entry in an index for a document structure definition to which said document complies, the entry comprising a mapping from said calculated number to said
15   document comprising the at least part of the branch.

11.   A method as claimed in claim 10, wherein creating an entry in the index comprises associating the document in the mapping to a value domain partition, the value domain partition comprising a value associated with the branch.

20

12.   A method as claimed in claim 10, wherein:

-          a further step comprises creating (57) an empty further index in which each integer from a further range of integers can be mapped to a document structure definition;

-          a further step comprises calculating (59) a further number for at least part of
25   said branch by applying a further hash function to said branch, the further number being limited to the further range of integers and the calculation possibly producing a same further number for different branches, and

-          a further step comprises creating (61) an entry in the further index, the entry in the further index comprising a mapping from the calculated further number to said document
30   structure definition to which said document complies.

13.   A computer program product enabling a programmable device to carry out a method as claimed in claim 10.

23

14.   An electronic device (71) for indexing a collection of documents, comprising electronic circuitry (73), the electronic circuitry (73) functionally comprising:

-         an index creator (75) for creating an empty index for each document structure definition of a collection of document structure definitions, the index mapping an integer from a range of integers to a document of the collection of documents;

-         a hash calculator (77) for calculating a number for at least a part of a branch in a document of the collection of documents by applying a hash function to the at least part of the branch, the number being limited to the range of integers and the calculation possibly producing a same number for different branches; and

-         an index filler (79) for creating an entry in an index for a document structure definition to which said document complies, the entry comprising a mapping from said calculated number to said document comprising the at least part of the branch.